SEQUENCE LISTING

(1) GENERAL INFORMATION:

    (i) APPLICANT: KIEFER, MICHAEL C.
                BARR, PHILIP J.

    (ii) TITLE OF INVENTION: NOVEL APOPTOSIS-MODULATING PROTEINS, DNA
           ENCODING THE PROTEINS AND METHODS OF USE THEREOF

    (iii) NUMBER OF SEQUENCES: 22

    (iv) CORRESPONDENCE ADDRESS:
        (A) ADDRESSEE: MORRISON & FOERSTER
        (B) STREET: 755 Page Mill Road
        (C) CITY: Palo Alto
        (D) STATE: California
        (E) COUNTRY: USA
        (F) ZIP: 94304-1018

    (v) COMPUTER READABLE FORM:
        (A) MEDIUM TYPE: Floppy disk
        (B) COMPUTER: IBM PC compatible
        (C) OPERATING SYSTEM: PC-DOS/MS-DOS
        (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

    (vi) CURRENT APPLICATION DATA:
        (A) APPLICATION NUMBER: US 08/320,157
        (B) FILING DATE: 07-OCT-1994
        (C) CLASSIFICATION:

    (viii) ATTORNEY/AGENT INFORMATION:
        (A) NAME: LEHNHARDT, SUSAN K.
        (B) REGISTRATION NUMBER: 33,943
        (C) REFERENCE/DOCKET NUMBER: 23647-20007.20

    (ix) TELECOMMUNICATION INFORMATION:
        (A) TELEPHONE: (415) 813-5600
        (B) TELEFAX: (415) 494-0792
        (C) TELEX: 706141

(2) INFORMATION FOR SEQ ID NO:1:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 8 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

    Asp Trp Gly Arg Val Val Ala Ile
    1   5

(2) INFORMATION FOR SEQ ID NO:2:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 36 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

```
(ix) FEATURE:
     (A) NAME/KEY: misc_difference
     (B) LOCATION: replace(23, "")
     (D) OTHER INFORMATION: /note= "This position is inosine."

(ix) FEATURE:
     (A) NAME/KEY: misc_difference
     (B) LOCATION: replace(27, "")
     (D) OTHER INFORMATION: /note= "This position is inosine."


(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
```

AGATCTGAAT TCAACTTGGG GGNCAGNAGT NGTNGC   36

```
(2) INFORMATION FOR SEQ ID NO:3:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 11 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear


    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

    Asp Trp Gly Gly Gln Glu Asn Asp Gln Ile Trp
    1               5                   10

(2) INFORMATION FOR SEQ ID NO:4:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 29 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

(ix) FEATURE:
     (A) NAME/KEY: misc_difference
     (B) LOCATION: replace(6, "")
     (D) OTHER INFORMATION: /note= "This position is inosine."

(ix) FEATURE:
     (A) NAME/KEY: misc_difference
     (B) LOCATION: replace(9, "")
     (D) OTHER INFORMATION: /note= "This position is inosine."


    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:
```

AGGGTNGGNG GNACNAGAGA CATCTAGGT   29

```
(2) INFORMATION FOR SEQ ID NO:5:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 41 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear


(ix) FEATURE:
     (A) NAME/KEY: misc_difference
     (B) LOCATION: replace(19, "")
     (D) OTHER INFORMATION: /note= "This position is inosine."
```

```
     (ix) FEATURE:
          (A) NAME/KEY: misc_difference
          (B) LOCATION: replace(22, "")
          (D) OTHER INFORMATION: /note= "This position is inosine."

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AGATCTAAGC TTGTCCCANC CNCCNTGNTC CTTGAGATCC A                    41

(2) INFORMATION FOR SEQ ID NO:6:

     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 2094 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear


     (ix) FEATURE:
          (A) NAME/KEY: CDS
          (B) LOCATION: 201..833

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GAGGATCTAC AGGGGACAAG TAAAGGCTAC ATCCAGATGC CGGGAATGCA CTGACGCCCA     60

TTCCTGGAAA CTGGGCTCCC ACTCAGCCCC TGGGAGCAGC AGCCGCCAGC CCCTCGGACC    120

TCCATCTCCA CCCTGCTGAG CCACCCGGGT TGGGCCAGGA TCCCGGCAGG CTGATCCCGT    180

CCTCCACTGA GACCTGAAAA ATG GCT TCG GGG CAA GGC CCA GGT CCT CCC       230
                        Met Ala Ser Gly Gln Gly Pro Gly Pro Pro
                         1               5                  10

AGG CAG GAG TGC GGA GAG CCT GCC CTG CCC TCT GCT TCT GAG GAG CAG     278
Arg Gln Glu Cys Gly Glu Pro Ala Leu Pro Ser Ala Ser Glu Glu Gln
             15                  20                  25

GTA GCC CAG GAC ACA GAG GAG GTT TTC CGC AGC TAC GTT TTT TAC CGC     326
Val Ala Gln Asp Thr Glu Glu Val Phe Arg Ser Tyr Val Phe Tyr Arg
             30                  35                  40

CAT CAG CAG GAA CAG GAG GCT GAA GGG GTG GCT GCC CCT GCC GAC CCA     374
His Gln Gln Glu Gln Glu Ala Glu Gly Val Ala Ala Pro Ala Asp Pro
             45                  50                  55

GAG ATG GTC ACC TTA CCT CTG CAA CCT AGC AGC ACC ATG GGG CAG GTG     422
Glu Met Val Thr Leu Pro Leu Gln Pro Ser Ser Thr Met Gly Gln Val
         60                  65                  70

GGA CGG CAG CTC GCC ATC ATC GGG GAC GAC ATC AAC CGA CGC TAT GAC     470
Gly Arg Gln Leu Ala Ile Ile Gly Asp Asp Ile Asn Arg Arg Tyr Asp
 75                  80                  85                  90

TCA GAG TTC CAG ACC ATG TTG CAG CAC CTG CAG CCC ACG GCA GAG AAT     518
Ser Glu Phe Gln Thr Met Leu Gln His Leu Gln Pro Thr Ala Glu Asn
             95                  100                 105

GCC TAT GAG TAC TTC ACC AAG ATT GCC ACC AGC CTG TTT GAG AGT GGC     566
Ala Tyr Glu Tyr Phe Thr Lys Ile Ala Thr Ser Leu Phe Glu Ser Gly
             110                 115                 120

ATC AAT TGG GGC CGT GTG GTG GCT CTT CTG GGC TTC GGC TAC CGT CTG     614
Ile Asn Trp Gly Arg Val Val Ala Leu Leu Gly Phe Gly Tyr Arg Leu
             125                 130                 135
```

```
GCC CTA CAC GTC TAC CAG CAT GGC CTG ACT GGC TTC CTA GGC CAG GTG        662
Ala Leu His Val Tyr Gln His Gly Leu Thr Gly Phe Leu Gly Gln Val
    140                 145                 150

ACC CGC TTC GTG GTC GAC TTC ATG CTG CAT CAC TGC ATT GCC CGG TGG        710
Thr Arg Phe Val Val Asp Phe Met Leu His His Cys Ile Ala Arg Trp
155                 160                 165                 170

ATT GCA CAG AGG GGT GGC TGG GTG GCA GCC CTG AAC TTG GGC AAT GGT        758
Ile Ala Gln Arg Gly Gly Trp Val Ala Ala Leu Asn Leu Gly Asn Gly
                175                 180                 185

CCC ATC CTG AAC GTG CTG GTG GTT CTG GGT GTG GTT CTG TTG GGC CAG        806
Pro Ile Leu Asn Val Leu Val Val Leu Gly Val Val Leu Leu Gly Gln
                190                 195                 200

TTT GTG GTA CGA AGA TTC TTC AAA TCA TGACTCCCAA GGGTGCCCTT             853
Phe Val Val Arg Arg Phe Phe Lys Ser
        205                 210
```

```
TGGGTCCCGG TTCAGACCCC TGCCTGGACT TAAGCGAAGT CTTTGCCTTC TCTGTTCCCT      913

TGCAGGGTCC CCCCTCAAGA GTACAGAAGC TTTAGCAAGT GTGCACTCCA GCTTCGGAGG      973

CCCTGCGTGG GGGCCAGTCA GGCTGCAGAG GCACCTCAAC ATTGCATGGT GCTAGTGCCC     1033

TCTCTCTGGG CCCAGGGCTG TGGCCGTCTC CTCCCTCAGC TCTCTGGGAC CTCCTTAGCC     1093

CTGTCTGCTA GGCGCTGGGG AGACTGATAA CTTGGGGAGG CAAGAGACTG GGAGCCACTT     1153

CTCCCCAGAA AGTGTTTAAC GGTTTTAGCT TTTTATAATA CCCTTGTGAG AGCCCATTCC     1213

CACCATTCTA CCTGAGGCCA GGACGTCTGG GGTGTGGGGA TTGGTGGGTC TATGTTCCCC     1273

AGGATTCAGC TATTCTGGAA GATCAGCACC CTAAGAGATG GGACTAGGAC CTGAGCCTGG     1333

TCCTGGCCGT CCCTAAGCAT GTGTCCCAGG AGCAGGACCT ACTAGGAGAG GGGGGCCAAG     1393

GTCCTGCTCA ACTCTACCCC TGCTCCCATT CCTCCCTCCG GCCATACTGC CTTTGCAGTT     1453

GGACTCTCAG GGATTCTGGG CTTGGGGTGT GGGGTGGGGT GGAGTCGCAG ACCAGAGCTG     1513

TCTGAACTCA CGTGTCAGAA GCCTCCAAGC CTGCCTCCCA AGGTCCTCTC AGTTCTCTCC     1573

CTTCCTCTCT CCTTATAGAC ACTTGCTCCC AACCCATTCA CTACAGGTGA AGGCTCTCAC     1633

CCATCCCTGG GGGCCTTGGG TGAGTGGCCT GCTAAGGCTC CTCCTTGCCC AGACTACAGG     1693

GCTTAGGACT TGGTTTGTTA TATCAGGGAA AAGGAGTAGG GAGTTCATCT GGAGGGTTCT     1753

AAGTGGGAGA AGGACTATCA ACACCACTAG GAATCCCAGA GGTGGATCCT CCCTCATGGC     1813

TCTGGCACAG TGTAATCCAG GGGTGTAGAT GGGGGAACTG TGAATACTTG AACTCTGTTC     1873

CCCCACCCTC CATGCTCCTC ACCTGTCTAG GTCTCCTCAG GGTGGGGGGT GACAGTGCCT     1933

TCTCTATTGG CACAGCCTAG GGTCTTGGGG GTCAGGGGGG AGAAGTTCTT GATTCAGCCA     1993

AATGCAGGGA GGGGAGGCAG ATGGAGCCCA TAGGCCACCC CCTATCCTCT GAGTGTTTGG     2053

AAATAAACTG TGCAATCCCC TCAAAAAAAA AACGGAGATC C                         2094
```

(2) INFORMATION FOR SEQ ID NO:7:

     (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 211 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

```
Met Ala Ser Gly Gln Gly Pro Gly Pro Pro Arg Gln Glu Cys Gly Glu
 1           5                   10                  15

Pro Ala Leu Pro Ser Ala Ser Glu Glu Gln Val Ala Gln Asp Thr Glu
            20              25                  30

Glu Val Phe Arg Ser Tyr Val Phe Tyr Arg His Gln Gln Glu Gln Glu
        35                  40                  45

Ala Glu Gly Val Ala Ala Pro Ala Asp Pro Glu Met Val Thr Leu Pro
    50                  55                  60

Leu Gln Pro Ser Ser Thr Met Gly Gln Val Gly Arg Gln Leu Ala Ile
 65              70                  75                      80

Ile Gly Asp Asp Ile Asn Arg Arg Tyr Asp Ser Glu Phe Gln Thr Met
            85                  90                  95

Leu Gln His Leu Gln Pro Thr Ala Glu Asn Ala Tyr Glu Tyr Phe Thr
            100                 105                 110

Lys Ile Ala Thr Ser Leu Phe Glu Ser Gly Ile Asn Trp Gly Arg Val
            115             120                 125

Val Ala Leu Leu Gly Phe Gly Tyr Arg Leu Ala Leu His Val Tyr Gln
    130                 135                 140

His Gly Leu Thr Gly Phe Leu Gly Gln Val Thr Arg Phe Val Val Asp
145                 150                 155                 160

Phe Met Leu His His Cys Ile Ala Arg Trp Ile Ala Gln Arg Gly Gly
                165             170                 175

Trp Val Ala Ala Leu Asn Leu Gly Asn Gly Pro Ile Leu Asn Val Leu
                180             185                 190

Val Val Leu Gly Val Val Leu Leu Gly Gln Phe Val Val Arg Arg Phe
        195             200                 205

Phe Lys Ser
210
```

(2) INFORMATION FOR SEQ ID NO:8:

     (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 1287 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ix) FEATURE:
        (A) NAME/KEY: CDS
        (B) LOCATION: 544..1176

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```
TTTTAATATA AATTAATGTG CTCTATTTAT AGAGACAATA CATGAAATAT ACTTAATAAA        60

AATTCAAATG TTATAGAACT GAAAAGATG AAAAGTAAAA ACAACCTATT CCCCAGAGGT         120

AGCCACTGTC CATAGTTTCT ATTTTAGATT CTTTCCTTTA TACAAGATTA TTATAGCTTC        180

TATTTTTTGG TGTATGAACT GTAGTCCTAG AGGATTTTAT TAGTTATGAG TTCTATAACT        240

AAGATCCATC ATCTTAGTTG CTAAGAACGT AGATACTGAG AACATCATTT AAAAAAACAT        300

TTTTGGCTGG CACCTCATGA TCACTGGAGT CTCGCGGGTC CCTCAGGCTG CACAGGGACA        360

AGTAAAGGCT ACATCCAGAT GCTGGGAATG CACTGACGCC CATTCCTGGA AACTGGGCTC        420

CCACTCAGCC CCTGGGAGCA GCAGCCGCCA GCCCCTCGGG ACCTCCATCT CCACCCTGCT        480

GAGCCACCCG GGTTGGGCCA GGATCCCGGC AGGCTGATCC CGTCCTCCAC TGAGACCTGA        540
```

```
AAA ATG GCT TCG GGG CAA GGC CCA GGT CCT CCC AGG CAG GAG TGC GGA        588
    Met Ala Ser Gly Gln Gly Pro Gly Pro Pro Arg Gln Glu Cys Gly
            215             220             225

GAG CCT GCC CTG CCC TCT GCT TCT GAG GAG CAG GTA GCC CAG GAC ACA        636
Glu Pro Ala Leu Pro Ser Ala Ser Glu Glu Gln Val Ala Gln Asp Thr
            230             235             240

GAG GAG GTT TTC CGC AGC TAC GTT TTT TAC CAC CAT CAG CAG GAA CAG        684
Glu Glu Val Phe Arg Ser Tyr Val Phe Tyr His His Gln Gln Glu Gln
            245             250             255

GAG GCT GAA GGG GCG GCT GCC CCT GCC GAC CCA GAG ATG GTC ACC TTA        732
Glu Ala Glu Gly Ala Ala Ala Pro Ala Asp Pro Glu Met Val Thr Leu
    260             265             270

CCT CTG CAA CCT AGC AGC ACC ATG GGG CAG GTG GGA CGG CAG CTC GCC        780
Pro Leu Gln Pro Ser Ser Thr Met Gly Gln Val Gly Arg Gln Leu Ala
275             280             285             290

ATC ATT GGG GAC GAC ATC AAC CGA CGC TAT GAC TCA GAG TTC CAG ACC        828
Ile Ile Gly Asp Asp Ile Asn Arg Arg Tyr Asp Ser Glu Phe Gln Thr
            295             300             305

ATG TTG CAG CAC CTG CAG CCC ACG GCA GAG AAT GCC TAT GAG TAC TTC        876
Met Leu Gln His Leu Gln Pro Thr Ala Glu Asn Ala Tyr Glu Tyr Phe
            310             315             320

ACC AAG ATT GCC TCC AGC CTG TTT GAG AGT GGC ATC AAT TGG GGC CGT        924
Thr Lys Ile Ala Ser Ser Leu Phe Glu Ser Gly Ile Asn Trp Gly Arg
            325             330             335

GTG GTG GCT CTT CTG GGC TTC AGC TAC CGT CTG GCC CTA CAC ATC TAC        972
Val Val Ala Leu Leu Gly Phe Ser Tyr Arg Leu Ala Leu His Ile Tyr
    340             345             350

CAG CGT GGC CTG ACT GGC TTC CTG GGC CAG GTG ACC CGC TTT GTG GTG        1020
Gln Arg Gly Leu Thr Gly Phe Leu Gly Gln Val Thr Arg Phe Val Val
355             360             365             370

GAC TTC ATG CTG CAT CAC TGC ATT GCC CGG TGG ATT GCA CAG AGG GGT        1068
Asp Phe Met Leu His His Cys Ile Ala Arg Trp Ile Ala Gln Arg Gly
            375             380             385
```

```
GGC TGG GTG GCA GCC CTG AAC TTG GGC AAT GGT CCC ATC CTG AAC GTG          1116
Gly Trp Val Ala Ala Leu Asn Leu Gly Asn Gly Pro Ile Leu Asn Val
            390                     395                 400

CTG GTG GTT CTG GGT GTG GTT CTG TTG GGC CAG TTT GTG GTA CGA AGA          1164
Leu Val Val Leu Gly Val Val Leu Leu Gly Gln Phe Val Val Arg Arg
            405                     410                 415

TTC TTC AAA TCA TGACTCCCAA GGGTGCCTTT GGGGTCCCAG TTCAGACCCC             1216
Phe Phe Lys Ser
            420

TGCCTGGACT TAAGCGAAGT CTTTGCCTTC TCTGCTCCTT GCAGGGTCCC CCCTCAAGAG       1276

TACAGAAGCT T                                                           1287
```

(2) INFORMATION FOR SEQ ID NO:9:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 211 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```
Met Ala Ser Gly Gln Gly Pro Gly Pro Pro Arg Gln Glu Cys Gly Glu
 1               5               10              15

Pro Ala Leu Pro Ser Ala Ser Glu Glu Gln Val Ala Gln Asp Thr Glu
            20              25              30

Glu Val Phe Arg Ser Tyr Val Phe Tyr His His Gln Gln Glu Gln Glu
            35              40              45

Ala Glu Gly Ala Ala Ala Pro Ala Asp Pro Glu Met Val Thr Leu Pro
    50              55              60

Leu Gln Pro Ser Ser Thr Met Gly Gln Val Gly Arg Gln Leu Ala Ile
65              70              75              80

Ile Gly Asp Asp Ile Asn Arg Arg Tyr Asp Ser Glu Phe Gln Thr Met
            85              90              95

Leu Gln His Leu Gln Pro Thr Ala Glu Asn Ala Tyr Glu Tyr Phe Thr
            100             105             110

Lys Ile Ala Ser Ser Leu Phe Glu Ser Gly Ile Asn Trp Gly Arg Val
            115             120             125

Val Ala Leu Leu Gly Phe Ser Tyr Arg Leu Ala Leu His Ile Tyr Gln
    130             135             140

Arg Gly Leu Thr Gly Phe Leu Gly Gln Val Thr Arg Phe Val Val Asp
145             150             155             160

Phe Met Leu His His Cys Ile Ala Arg Trp Ile Ala Gln Arg Gly Gly
            165             170             175

Trp Val Ala Ala Leu Asn Leu Gly Asn Gly Pro Ile Leu Asn Val Leu
            180             185             190
```

```
Val Val Leu Gly Val Val Leu Leu Gly Gln Phe Val Val Arg Arg Phe
        195                 200                 205

Phe Lys Ser
    210
```

(2) INFORMATION FOR SEQ ID NO:10:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 211 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```
Met Ala Ser Gly Gln Gly Pro Gly Pro Pro Arg Gln Glu Cys Gly Glu
1               5                   10                  15

Pro Ala Leu Pro Ser Ala Ser Glu Glu Gln Val Ala Gln Asp Thr Glu
            20                  25                  30

Glu Val Phe Arg Ser Tyr Val Phe Tyr Arg His Gln Gln Glu Gln Glu
            35                  40                  45

Ala Glu Gly Val Ala Ala Pro Ala Asp Pro Glu Met Val Thr Leu Pro
    50                  55                  60

Leu Gln Pro Ser Ser Thr Met Gly Gln Val Gly Arg Gln Leu Ala Ile
65                  70                  75                  80

Ile Gly Asp Asp Ile Asn Arg Arg Tyr Asp Ser Glu Phe Gln Thr Met
                85                  90                  95

Leu Gln His Leu Gln Pro Thr Ala Glu Asn Ala Tyr Glu Tyr Phe Thr
            100                 105                 110

Lys Ile Ala Thr Ser Leu Phe Glu Ser Gly Ile Asn Trp Gly Arg Val
        115                 120                 125

Val Ala Leu Leu Gly Phe Gly Tyr Arg Leu Ala Leu His Val Tyr Gln
    130                 135                 140

His Gly Leu Thr Gly Phe Leu Gly Gln Val Thr Arg Phe Val Val Asp
145                 150                 155                 160

Phe Met Leu His His Cys Ile Ala Arg Trp Ile Ala Gln Arg Gly Gly
                165                 170                 175

Trp Val Ala Ala Leu Asn Leu Gly Asn Gly Pro Ile Leu Asn Val Leu
            180                 185                 190

Val Val Leu Gly Val Val Leu Leu Gly Gln Phe Val Val Arg Arg Phe
        195                 200                 205

Phe Lys Ser
    210
```

(2) INFORMATION FOR SEQ ID NO:11:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 211 amino acids
        (B) TYPE: amino acid

        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

    Met Ala Ser Gly Gln Gly Pro Gly Pro Pro Arg Gln Glu Cys Gly Glu
    1   5   10   15

    Pro Ala Leu Pro Ser Ala Ser Glu Glu Gln Val Ala Gln Asp Thr Glu
                20   25   30

    Glu Val Phe Arg Ser Tyr Val Phe Tyr His His Gln Gln Glu Gln Glu
            35   40   45

    Ala Glu Gly Ala Ala Ala Pro Ala Asp Pro Glu Met Val Thr Leu Pro
        50   55   60

    Leu Gln Pro Ser Ser Thr Met Gly Gln Val Gly Arg Gln Leu Ala Ile
    65   70   75   80

    Ile Gly Asp Asp Ile Asn Arg Arg Tyr Asp Ser Glu Phe Gln Thr Met
                85   90   95

    Leu Gln His Leu Gln Pro Thr Ala Glu Asn Ala Tyr Glu Tyr Phe Thr
                100   105   110

    Lys Ile Ala Ser Ser Leu Phe Glu Ser Gly Ile Asn Trp Gly Arg Val
                115   120   125

    Val Ala Leu Leu Gly Phe Ser Tyr Arg Leu Ala Leu His Ile Tyr Gln
        130   135   140

    Arg Gly Leu Thr Gly Phe Leu Gly Gln Val Thr Arg Phe Val Val Asp
    145   150   155   160

    Phe Met Leu His His Cys Ile Ala Arg Trp Ile Ala Gln Arg Gly Gly
                165   170   175

    Trp Val Ala Ala Leu Asn Leu Gly Asn Gly Pro Ile Leu Asn Val Leu
                180   185   190

    Val Val Leu Gly Val Val Leu Leu Gly Gln Phe Val Val Arg Arg Phe
        195   200   205

    Phe Lys Ser
    210

(2) INFORMATION FOR SEQ ID NO:12:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 239 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

    Met Ala His Ala Gly Arg Thr Gly Tyr Asp Asn Arg Glu Ile Val Met
    1   5   10   15

    Lys Tyr Ile His Tyr Lys Leu Ser Gln Arg Gly Tyr Glu Trp Asp Ala
                20   25   30

```
Gly Asp Val Gly Ala Ala Pro Pro Gly Ala Ala Pro Ala Pro Gly Ile
        35              40                  45

Phe Ser Ser Gln Pro Gly His Thr Pro His Thr Ala Ala Ser Arg Asp
    50              55                  60

Pro Val Ala Arg Thr Ser Pro Leu Gln Thr Pro Ala Ala Pro Gly Ala
65              70                  75                  80

Ala Ala Gly Pro Ala Leu Ser Pro Val Pro Pro Val Val His Leu Thr
            85              90                  95

Leu Arg Gln Ala Gly Asp Asp Phe Ser Arg Arg Tyr Arg Arg Asp Phe
            100             105             110

Ala Glu Met Ser Arg Gln Leu His Leu Thr Pro Phe Thr Ala Arg Gly
        115             120             125

Arg Phe Ala Thr Val Val Glu Glu Leu Phe Arg Asp Gly Val Asn Trp
    130             135             140

Gly Arg Ile Val Ala Phe Phe Glu Phe Gly Gly Val Met Cys Val Glu
145             150             155             160

Ser Val Asn Arg Glu Met Ser Pro Leu Val Asp Asn Ile Ala Leu Trp
            165             170             175

Met Thr Glu Tyr Leu Asn Arg His Leu His Thr Trp Ile Gln Asp Asn
            180             185             190

Gly Gly Trp Asp Ala Phe Val Glu Leu Tyr Gly Pro Ser Met Arg Pro
            195             200             205

Leu Phe Asp Phe Ser Trp Leu Ser Leu Lys Thr Leu Leu Ser Leu Ala
    210             215             220

Leu Val Gly Ala Cys Ile Thr Leu Gly Ala Tyr Leu Gly His Lys
225             230             235
```

(2) INFORMATION FOR SEQ ID NO:13:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 192 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

```
Met Asp Gly Ser Gly Glu Gln Pro Arg Gly Gly Gly Pro Thr Ser Ser
1               5                   10                  15

Glu Gln Ile Met Lys Thr Gly Ala Leu Leu Leu Gln Gly Phe Ile Gln
            20                  25                  30

Asp Arg Ala Gly Arg Met Gly Gly Glu Ala Pro Glu Leu Ala Leu Asp
        35              40                  45

Pro Val Pro Gln Asp Ala Ser Thr Lys Lys Leu Ser Glu Cys Leu Lys
    50              55                  60

Arg Ile Gly Asp Glu Leu Asp Ser Asn Met Glu Leu Gln Arg Met Ile
65              70                  75                  80
```

```
Ala Ala Val Asp Thr Asp Ser Pro Arg Glu Val Phe Phe Arg Val Ala
              85              90                  95

Ala Asp Met Phe Ser Asp Gly Asn Phe Asn Trp Gly Arg Val Val Ala
             100              105             110

Leu Phe Tyr Phe Ala Ser Lys Leu Val Leu Lys Ala Leu Cys Thr Lys
             115              120             125

Val Pro Glu Leu Ile Arg Thr Ile Met Gly Trp Thr Leu Asp Phe Leu
         130              135             140

Arg Glu Arg Leu Leu Gly Trp Ile Gln Asp Gln Gly Gly Trp Asp Gly
145              150              155             160

Leu Leu Ser Tyr Phe Gly Thr Pro Thr Trp Gln Thr Val Thr Ile Phe
             165              170             175

Val Ala Gly Val Leu Thr Ala Ser Leu Thr Ile Trp Lys Lys Met Gly
             180              185             190
```

(2) INFORMATION FOR SEQ ID NO:14:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 233 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

```
Met Ser Gln Ser Asn Arg Glu Leu Val Val Asp Phe Leu Ser Tyr Lys
1               5               10                  15

Leu Ser Gln Lys Gly Tyr Ser Trp Ser Gln Phe Ser Asp Val Glu Glu
             20              25              30

Asn Arg Thr Glu Ala Pro Glu Gly Thr Glu Ser Glu Met Glu Thr Pro
         35              40              45

Ser Ala Ile Asn Gly Asn Pro Ser Trp His Leu Ala Asp Ser Pro Ala
         50              55              60

Val Asn Gly Ala Thr Gly His Ser Ser Ser Leu Asp Ala Arg Glu Val
65              70              75                  80

Ile Pro Met Ala Ala Val Lys Gln Ala Leu Arg Glu Ala Gly Asp Glu
             85              90                  95

Phe Glu Leu Arg Tyr Arg Arg Ala Phe Ser Asp Leu Thr Ser Gln Leu
             100             105             110

His Ile Thr Pro Gly Thr Ala Tyr Gln Ser Phe Glu Gln Val Val Asn
             115             120             125

Glu Leu Phe Arg Asp Gly Val Asn Trp Gly Arg Ile Val Ala Phe Phe
         130             135             140

Ser Phe Gly Gly Ala Leu Cys Val Glu Ser Val Asp Lys Glu Met Gln
145             150             155             160

Val Leu Val Ser Arg Ile Ala Ala Trp Met Ala Thr Tyr Leu Asn Asp
             165             170             175
```

His Leu Glu Pro Trp Ile Gln Glu Asn Gly Gly Trp Asp Thr Phe Val
           180   185   190

Glu Leu Tyr Gly Asn Asn Ala Ala Ala Glu Ser Arg Lys Gly Gln Glu
        195   200   205

Arg Phe Asn Arg Trp Phe Leu Thr Gly Met Thr Val Ala Gly Val Val
        210   215   220

Leu Leu Gly Ser Leu Phe Ser Arg Lys
225   230

(2) INFORMATION FOR SEQ ID NO:15:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 226 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Glu Leu Asp Gly Tyr Glu Pro Glu Pro Leu Gly Lys Arg Pro Ala Val
1   5   10   15

Leu Pro Leu Leu Glu Leu Val Gly Glu Ser Gly Asn Asn Thr Ser Thr
        20   25   30

Asp Gly Ser Leu Pro Ser Thr Pro Pro Pro Ala Glu Glu Glu Glu Asp
        35   40   45

Glu Leu Tyr Arg Gln Ser Leu Glu Ile Ile Ser Arg Tyr Leu Arg Glu
        50   55   60

Gln Ala Thr Gly Ala Lys Asp Thr Lys Pro Met Gly Arg Ser Gly Ala
65   70   75   80

Thr Ser Arg Lys Ala Leu Glu Thr Leu Arg Arg Val Gly Asp Gly Val
          85   90   95

Gln Arg Asn His Glu Thr Val Phe Gln Gly Met Leu Arg Lys Leu Asp
        100   105   110

Ile Lys Asn Glu Asp Asp Val Lys Ser Leu Ser Arg Val Met Ile His
        115   120   125

Val Phe Ser Asp Gly Val Thr Asn Trp Gly Arg Ile Val Thr Leu Ile
        130   135   140

Ser Phe Gly Ala Phe Val Ala Lys His Leu Lys Thr Ile Asn Gln Glu
145   150   155   160

Ser Cys Ile Glu Pro Leu Ala Glu Ser Ile Thr Asp Val Leu Val Arg
          165   170   175

Thr Lys Arg Asp Trp Leu Val Lys Gln Arg Gly Trp Asp Gly Phe Val
          180   185   190

Glu Phe Phe His Val Glu Asp Leu Glu Gly Gly Ile Arg Asn Val Leu
        195   200   205

```
Leu Ala Phe Ala Gly Val Ala Gly Val Gly Ala Gly Leu Ala Tyr Leu
    210             215             220

Ile Arg
225
```

(2) INFORMATION FOR SEQ ID NO:16:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 172 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

```
Met Ala Glu Ser Glu Leu Met His Ile His Ser Leu Ala Glu His Tyr
1           5               10              15

Leu Gln Tyr Val Leu Gln Val Pro Ala Phe Glu Ser Ala Pro Ser Gln
        20              25              30

Ala Cys Arg Val Leu Gln Arg Val Ala Phe Ser Val Gln Lys Glu Val
        35              40              45

Glu Lys Asn Leu Lys Ser Tyr Leu Asp Asp Phe His Val Glu Ser Ile
    50              55              60

Asp Thr Ala Arg Ile Ile Phe Asn Gln Val Met Glu Lys Glu Phe Glu
65              70              75              80

Asp Gly Ile Ile Asn Trp Gly Arg Ile Val Thr Ile Phe Ala Phe Gly
                85              90              95

Gly Val Leu Leu Lys Lys Leu Pro Gln Glu Gln Ile Ala Leu Asp Val
            100             105             110

Cys Ala Tyr Lys Gln Val Ser Ser Phe Val Ala Glu Phe Ile Met Asn
        115             120             125

Asn Thr Gly Glu Trp Ile Arg Gln Asn Gly Gly Trp Glu Asp Gly Phe
    130             135             140

Ile Lys Lys Phe Glu Pro Lys Ser Gly Trp Leu Thr Phe Leu Gln Met
145             150             155             160

Thr Gly Gln Ile Trp Glu Met Leu Phe Leu Leu Lys
                165             170
```

(2) INFORMATION FOR SEQ ID NO:17:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 187 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

```
Met Ala Tyr Ser Thr Arg Glu Ile Leu Leu Ala Leu Cys Ile Arg Asp
1           5               10              15
```

Ser Arg Val His Gly Asn Gly Thr Leu His Pro Val Leu Glu Leu Ala
20 25 30

Ala Arg Glu Thr Pro Leu Arg Leu Ser Pro Glu Asp Thr Val Val Leu
35 40 45

Arg Tyr His Val Leu Leu Glu Glu Ile Ile Glu Arg Asn Ser Glu Thr
50 55 60

Phe Thr Glu Thr Trp Asn Arg Phe Ile Thr His Thr Glu His Val Asp
65 70 75 80

Leu Asp Phe Asn Ser Val Phe Leu Glu Ile Phe His Asp Leu Ile Asn
85 90 95

Trp Gly Arg Ile Cys Gly Phe Ile Val Phe Ser Ala Arg Met Ala Lys
100 105 110

Tyr Cys Lys Asp Ala Asn Asn His Leu Glu Ser Thr Val Ile Thr Thr
115 120 125

Ala Tyr Asn Phe Ser Glu Gly Leu Asp Gly Trp Ile His Gln Gln Gly
130 135 140

Gly Trp Ser Thr Leu Ile Glu Asp Asn Ile Pro Gly Ser Arg Arg Phe
145 150 155 160

Ser Trp Thr Leu Phe Leu Ala Gly Leu Thr Leu Ser Leu Leu Val Ile
165 170 175

Cys Ser Tyr Leu Phe Ile Ser Arg Gly Arg His
180 185

(2) INFORMATION FOR SEQ ID NO:18:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 181 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Met Glu Gly Glu Glu Leu Ile Tyr His Asn Ile Ile Asn Glu Ile Leu
1 5 10 15

Val Gly Tyr Ile Lys Tyr Tyr Met Asn Asp Ile His Glu Leu Ser Pro
20 25 30

Tyr Gln Gln Gln Ile Lys Lys Ile Leu Thr Tyr Tyr Asp Glu Cys Leu
35 40 45

Asn Lys Gln Val Thr Ile Thr Phe Ser Leu Thr Asn Ala Gln Glu Ile
50 55 60

Lys Thr Gln Phe Thr Gly Val Val Thr Glu Leu Phe Lys Arg Gly Asp
65 70 75 80

Pro Ser Leu Gly Arg Ala Leu Ala Trp Met Ala Trp Cys Met His Ala
85 90 95

Cys Arg Thr Leu Cys Cys Asn Gln Ser Thr Pro Tyr Tyr Val Val Asp
100 105 110

```
Leu Ser Val Arg Gly Met Leu Glu Ala Met Lys His Asn Leu Leu Pro
        115                 120             125
```

```
Trp Met Ile Ser His Gly Gly Gln Glu Glu Phe Leu Ala Phe Ser Leu
        130             135                 140
```

```
His Ser Gln Ile Tyr Ser Val Ile Phe Asn Ile Lys Tyr Phe Leu Ser
145                 150                 155                 160
```

```
Lys Phe Cys Asn His His Phe Leu Arg Ser Cys Val Gln Leu Leu Arg
                165             170                 175
```

```
Lys Cys Asn Leu Ile
                180
```

(2) INFORMATION FOR SEQ ID NO:19:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 280 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

```
Met Thr Arg Cys Thr Ala Asp Asn Ser Leu Thr Asn Pro Ala Tyr Arg
1           5                   10                  15
```

```
Arg Arg Thr Met Ala Thr Gly Glu Met Lys Glu Phe Leu Gly Ile Lys
            20              25              30
```

```
Gly Thr Glu Pro Thr Asp Phe Gly Ile Asn Ser Asp Ala Gln Asp Leu
        35              40              45
```

```
Pro Ser Pro Ser Arg Gln Ala Ser Thr Arg Arg Met Ser Ile Gly Glu
    50              55              60
```

```
Ser Ile Asp Gly Lys Ile Asn Asp Trp Glu Glu Pro Arg Leu Asp Ile
65              70              75                  80
```

```
Glu Gly Phe Val Val Asp Tyr Phe Thr His Arg Ile Arg Gln Asn Gly
            85              90              95
```

```
Met Glu Trp Phe Gly Ala Pro Gly Leu Pro Cys Gly Val Gln Pro Glu
        100             105             110
```

```
His Glu Met Met Arg Val Met Gly Thr Ile Phe Glu Lys Lys His Ala
        115             120             125
```

```
Glu Asn Phe Glu Thr Phe Cys Glu Gln Leu Leu Ala Val Pro Arg Ile
    130             135             140
```

```
Ser Phe Ser Leu Tyr Gln Asp Val Val Arg Thr Val Gly Asn Ala Gln
145             150             155             160
```

```
Thr Asp Gln Cys Pro Met Ser Tyr Gly Arg Leu Ile Gly Leu Ile Ser
            165             170             175
```

```
Phe Gly Gly Phe Val Ala Ala Lys Met Met Glu Ser Val Glu Leu Gln
        180             185             190
```

```
Gly Gln Val Arg Asn Leu Phe Val Tyr Thr Ser Leu Phe Ile Lys Thr
    195             200             205
```

```
    Arg Ile Arg Asn Asn Trp Lys Glu His Asn Arg Ser Trp Asp Asp Phe
        210             215             220

    Met Thr Leu Gly Lys Gln Met Lys Glu Asp Tyr Glu Arg Ala Glu Ala
    225             230             235             240

    Glu Lys Val Gly Arg Arg Lys Gln Asn Arg Arg Trp Ser Met Ile Gly
                245             250             255

    Ala Gly Val Thr Ala Gly Ala Ile Gly Ile Val Gly Val Val Val Cys
                260             265             270

    Gly Arg Met Met Phe Ser Leu Lys
                275             280
```

(2) INFORMATION FOR SEQ ID NO:20:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 5408 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ix) FEATURE:
        (A) NAME/KEY: CDS
        (B) LOCATION: 1665..1928

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

```
GAATTCTGGT AATTAGTTAA ACAACCTTGA ACAAGTTGTT TCACTTCTCT GAGTCTCAGT        60

TTCTCACTCA AAAATGGTGA ATAATTTGTA AGACTTCGCT AATAATCTAC GACTCTACAA       120

GAGGCAATAG GGTACTGTGG ACAGAGAGCA GGCTTTGGAA ACACACAAGA CTGGGTTTAG       180

ATTCCTGCAC TCCACCCAGT GTGTGACTTG GCCAAGCTTC TTCACTTCTC TAAACCCCCA       240

TCTGTGTATC TGTACAGGAA TGAATGAATG AGTATGTGCA GCCAAGCTAT GCAAACTCCA       300

GGTTAAAATA TTGCCTTGGG TTTTTTAGTA AATTGTTCAA GCCCATGACA TTCTAGCAGA       360

AAAAGCCTAG TGTCTCTTTC TTAAGGTGAT TGTGTCCATG TGTTTTCCAG GAACTCTATG       420

GGTTTCTCAA CCCAAATTCA CCCTGCCCTT GACCAAATGG CTCACCAGCT TCACGGATGC       480

TGCTCTGATG ACACACCCTG CAGTCAGCAT CTGCCCCTGC AGCTAGAATG GATTTCTGAG       540

TGGGCATTAG CTGGGGGATA CCACATGGGC ACCAATGTCA CAGATCTTCT GTCACAGTCC       600

ACCCCGAACC ATTGCTTCTC AAATCATAAT CCCTTAGCAG GACAGCTAGG TGCAGCACGC       660

ATGACACAAA CACCAGCCCT TGCCTACAAT CTCAGCCACT ATCTTGAGTC TGAGCAACTA       720

GTCTAGTGGC AGCCGCGCCC TTCCTTTTCA AGAGAGTTCT GGGATCAGAT CCTTTCACAA       780

ACAGATCCCT CCCCACCCTG CCTGTTGTCC AGGTCTGCAC ACTGAAAAGT AAGACAGCAT       840

TTGCTAAGCC ATATTTCAAA AAGTTTGCTT ATACCTTCAT CTCAGGACAA CAAGTGCCTG       900

CTTAAGAGCC TTATGTTTGT GTAACTGGTA TTTTTTTTTC CCCTGACCTT CCAAGGCCTA       960

GTCTACTTTC TCCCTCCCTA GCTGAACAAA AGTGAAGTTG AAATAATTTG AACTACCCCT      1020
```

```
TTTAGTGGGC AGCCCATTTG ATTTTTACCT TAGCCAGAGC CTTAATTTGT CCATGTGAGC    1080

ATAGCAGTAC CTTGCAGCAC CTGAGGCACA ATACATTGTT TAAAGAGTGA CAGTGCGTCC    1140

CATTCCAATA AGAACCACAC TCAGAGCAAA GGTTCCCTCT CCTGTGTGGA GAGTGACCCA    1200

TGGTAGAAAA TTTGCAGACT TCGTTACCTC TTCATCAGTT GAAAAATCTA TTTATTCATT    1260

TATGCATTTA ATTTTCCCTA TCTAAGCCAG GGATAGTCAA ACATTTTCTG TAAAGGGCCA    1320

AGTAGCATGA TAAATATGTT AGGCTCTGCA GGCCACTTAC AGTTTTGTCA TGTATTCTTT    1380

TTTTGCTCCC TGTTTGTATT ATTTTGTTTA CAATGCTTTA AAAATGTAAA AAAACAGATG    1440

ATCACTGGAG TCTCACGGGT CCCTCGGGCC ACACAGGGAC AAGCAAAGGC TACATCCAGA    1500

TACCAGAAAT GCACTGACGC CCGTTCCTGG AAGCTGGGCT CCCACTCAGC CCCTGGGAGC    1560

AGCAGCCTCC AGCCCCTTGG GACCTTCAAC TCCACCCTGC TGACCCACGC GGGTTGAGCC    1620

AGCATCCCTG GAGGCTGACA CTGTCCTCCA CTGAGACCTG AAAA ATG GCA TCG GGG    1676
                                                Met Ala Ser Gly
                                                215
```

```
CAA GGC CCA GGG CCT CCC AGG CAG GAG TGC GGA AAG CCT GCC CTG CCC    1724
Gln Gly Pro Gly Pro Pro Arg Gln Glu Cys Gly Lys Pro Ala Leu Pro
        220             225                 230

TCT GCT TCT GAG GAG CAG GTA GCC CAG GAC ATG GAG GGG TTT TCC GCA    1772
Ser Ala Ser Glu Glu Gln Val Ala Gln Asp Met Glu Gly Phe Ser Ala
        235             240                 245

GCT ACG TTT TTT ACC ACC ATC AGC AGG AAC AGG AGG CTG AAG GGG CGG    1820
Ala Thr Phe Phe Thr Thr Ile Ser Arg Asn Arg Arg Leu Lys Gly Arg
        250             255                 260

CCG CCC CTG CCG ACC CAG AGA TGG TCA CCT TGC CCC TCC AAC CTA GCA    1868
Pro Pro Leu Pro Thr Gln Arg Trp Ser Pro Cys Pro Ser Asn Leu Ala
        265             270                 275

GCA CCA TGG GGC AGG TGG GAC GGC AGC TCG CCA TCA CCA GGA CGA CAT    1916
Ala Pro Trp Gly Arg Trp Asp Gly Ser Ser Pro Ser Pro Gly Arg His
280             285                 290                 295

CAA CCG GCA CTA TGACTTCGGA GTTCCAGACC ATGCTGCAGC ACCTGCAGCC    1968
Gln Pro Ala Leu
```

```
CACGGCAGAG AACGCCTACG AGTACTTCAC CAAGATCGCC TCCAGCCTGT TTGAGAGTGG    2028

CATCAACCGG GGCCGTGTGG TGGCTCTCCT GGGCTTCGGC TACCGTCTGG TCCTACATGT    2088

CTACCAGCAC GGCTTGACTG GCTTCCTGGG CCTGGTGACC CGCTTCGTGG TCTTCATGCT    2148

GCAACAAGGC ATCGCCCGGT GGATCTCGCA GAGGGGCGGC TGGGTGGCAG CCCTGGACTT    2208

GGGCAATAGT CCCATCCTGA CGTGCTGGT GGTTGTGGGT GTGGTTCTGC TGGGCCAGTT    2268

TGTGGTAAGA AGATTCTTCA AATCATGACT CCCAGGGGTG TCCTTTGGGG TCCCAGCTGT    2328

GACCCCTGCC TGGACTTAAG CCAAGTCTTT GCCTTCCCCA CTCCCTTGCA GGGGTCACCC    2388

TTCAAAAGTA CAGAAGCTCT AGCAAGTGTG CACCCCCGCT GCGGAGGGCC CCTGCGTGGG    2448

GGCCAGTCAG GCTGCGGAGG CACCTCAACA TTGCACGGTG CTAGTGGGCC CTCTCTCTGG    2508
```

```
GCCCAGGGGC TGTGCCCTCC TCCCTTGGCT CTCTGGGACC TCCTTAGTCT TGTCTGCTAG    2568

GCGCTGCAGA GGCTGATAAC TTGGGGAAGC AAGAGACTGG GAGCCACTCC TCCCCAGTAA    2628

GTGTTTAACG GTTTTAGCTT TTTATAATAC CCTTGGGAGA GCCCATTCCC ACCATTCTAC    2688

CCAAGGCCGG GATGTCTGGG GTGTGGGGGT TGGTGGGTCG TAACCTACGT GCCCCAGGAT    2748

TCAGCTATTC TGGAAGATCA GAGCCTAAGA GCTAGGACTT GATCCTGGTC CTGGCCGTCC    2808

CTAAGCATCA TGTGTCCCAG GAGCAGGACT GACTGGGAGA GGGGACCAAG GTCCTACCCA    2868

GCTCTCCCCG TGCCCCCATT CCTCCTCCGG CCATACTGCC TTTGCAGTTG GACTCTCAGG    2928

GATTCTGGGC TTGGGGTGTG GGGCGGCGTG GAGTAACAGG CCAGAGCTGT CTGAACTTAT    2988

GTGTCAGAAG CCTCCAAGCC TGCCTCCCAA GGTCCTCTCA GCTCTCTCCC TTCCTCTCTC    3048

CTTATAGATA CTTGCTCCCA ACCCATTCAC TACAGGTGAA GGCCCTCACC CATCCCTGGG    3108

GGCCTTGGGT GAGTGATGCG CTAAGGCCCC TCCCCGCCCA GACTACAGGG CTTGGTTTAG    3168

GGCTTGGTTT GTTATTTCAG GGATAAGGAG TAGGGAGTTC ATCTGGAAGG TTCTAAGTGG    3228

GAGAAGGACT ATCAACACCA CAGGAATCCC AGAGGTGGGA TCCTCCCTCA TGGCTCTGGC    3288

ACAGTGTAAT CCAGGGGTGG AGATAGGGAA CTGTGAATAC CTGAACTCTG TCCCCCGACC    3348

CTCCATGCTC CTCACCTTTC TGGGTCTCTC CTCAGTGTGG GGGTGAGAGT ACCTTCTCTA    3408

TCGGGCACAG CCTAGGGTGT TGGGGGTGAA GGGGGAGAAG TTCTTGATTC AGCCAAATGC    3468

AGGGAGGGGA GGCAGAAGGA GCCCACAGGC CACTCCCTAT CCTCTGAGTG TTTGGAAATA    3528

AACTGTGCAA TCCCATCAAA AAAAAAAAGG AGAAAAAAAT GTAAAAAACA TTCTTAGCTG    3588

TAAGCTACTT ATAGGGGGAT AAAGACAGGA CTGTTAATGG ACACAAACAT ACAGTTAGAG    3648

AGAAGAAATA AGTTCTGTCC AGGCACGGTG GCTCACACCT CTAACTCCAG CACTTTGGGA    3708

GACCAAAGTG GGAAGATCAT TTGAGTCCAG GAGTTCGAGA CCAGCCTGGA CAACATAGCA    3768

AGATCTTATC TCTACAGAAA ATTTAAAAAA AAGAAAAAAA CTAGCCGCAC AGGTCTGCAG    3828

TCCTAGCTAC TCGGGAGGCT AAGGTGGGAG AATCCTTGAA CCCAGGGATT TAGTTTGAGG    3888

TTGCAGTGAG CTATGATTGC ACCACTGCAC TCCAGACTGG GTGACTGAGT GAGACCCTGT    3948

CTCAAATATA AAGAAGGAAC AAGTTCTAGT TTTCAATAGC GCAATAGGGT GAGTGCAGTT    4008

AGCAACAACA TATTGTGTAT TTCAAAATAG CTACAAGAGA GGATATGAAG TGTTCCCCCA    4068

AACAAGGAAT GATAACGTTC GAGGTGACAG ATACCTTAAA TACCCTGATT TGATCATTAC    4128

ACATTCAATG TATGTATCAA AATATTACAT GTACCCCACA AATTTGTGTA AATATTATGT    4188

ATCCACTTTT TAAAGTTGGC AGAGCCCAAA AGCACTACTA TGGCTTCCAG TGGTCACTGT    4248

GAGCACTGCC AGCTCAGCAA ATGTATCACC CAAAATCTGG GCAATGTGGG AAATTGGCTT    4308

CATGGCAGCT ATGGCTTTGC CACTGATAGG AATGATTTCC AGAGATACTT AATCCTCAAT    4368

TCGGGACTCT TTGCTTCAGG AGTTTGGCTG GCCAGGAACA TGAGTGACAG TGACCTCTTG    4428

GCACTTCAGC TGGGGGTGTA GCCAAGCAGA CAAATGGAAT CTTGTGCTGA ACCCAAACCT    4488
```

```
TCTAGAAACA GAGCCTGTGA GCATCACAAG ATATGCCCTG ATGGAAGCTG AAGTTTAATT    4548

CAGCTGAGCG CTTGCCCCTT TCCAACCTGG TTTCTTTTTG TTCCTTGAGT CCAGTCAGAA    4608

TGCCATTCCC TGGCCAGCAG CCAGCCTTTA GTGACTGTCT CTGTTCTGCA AAGCTCTGTA    4668

TATAGTTACT GAGTTTCTGC AGGGGGTGAT CTTTGCTCTT GTCCTAAGAA ATAACTACAG    4728

TGTTTTAAGA AATATTTGAG GCCGGGTGCA GTGGTTCACA CCTGTAATCC AGCACTTTGG    4788

GAGGCCAAGG CAGGTGGATC ATGAGGTCAA GAGTTTGAGA CCATCATGGC CAACATGGTG    4848

AAACCCCATC TCTACTAAAA ATACAAAAAT TAGCTGGGTG TGGTGGCGGG CACCTGTAGT    4908

CCCAGCTACT CGGGAGGCTG AGGCAGGAGA ATCGCTTGAG CCTGGGAGGC GGAGGTTGCA    4968

CTGAGCCGAT ATCACGCCAC TGCACTCCAG CCTGGCGACA GAGCGAGACT CCATCTCAAA    5028

AAAAAGAAAA AATAAATAGT TGAAATAAAG ACTGCACATA AAGACAAAAA AAAAGTTTAT    5088

AAAGTTAAAA AATAAAATAA AAAACAGGCT CCAGGCTGGA TTGGGCCCAG AGGCTGTAGG    5148

ACACAGACCC CCAGCCAATG ACTTCATAAA TCCGGATGTT AATCAGCCTC ACCTGGGAAT    5208

TTGGGGAGGG GACTCATTTT AAAACAGTTT CCTGGATTCT AACCCAACCC AGAAAATCAG    5268

ACTCTTTGAG CTAAATTCTT AAGCTCCCTG GTGATGATGA TGGAACCAGT TTATGGCTGA    5328

CCCCAGAGTA CCGTCTGAAA GACGTGCCAC ATCCCTCTCT CTCCAGCCTC CCCTTCTCCT    5388

CCATTCCCCA GGGAGAATTC                                                5408
```

(2) INFORMATION FOR SEQ ID NO:21:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 88 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

```
Met Ala Ser Gly Gln Gly Pro Gly Pro Pro Arg Gln Glu Cys Gly Lys
1               5                  10                  15

Pro Ala Leu Pro Ser Ala Ser Glu Glu Gln Val Ala Gln Asp Met Glu
            20                  25                  30

Gly Phe Ser Ala Ala Thr Phe Phe Thr Thr Ile Ser Arg Asn Arg Arg
            35                  40                  45

Leu Lys Gly Arg Pro Pro Leu Pro Thr Gln Arg Trp Ser Pro Cys Pro
    50                  55                  60

Ser Asn Leu Ala Ala Pro Trp Gly Arg Trp Asp Gly Ser Ser Pro Ser
65                  70                  75                  80

Pro Gly Arg His Gln Pro Ala Leu
                85
```

(2) INFORMATION FOR SEQ ID NO:22:

    (i) SEQUENCE CHARACTERISTICS:

    (A) LENGTH: 210 amino acids
    (B) TYPE: amino acid
    (C) STRANDEDNESS: single
    (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Met Ala Ser Gly Gln Gly Pro Gly Pro Pro Arg Gln Glu Cys Gly Glu
1   5   10   15

Pro Ala Leu Pro Ser Ala Ser Glu Glu Gln Val Ala Gln Asp Thr Glu
            20   25   30

Glu Val Phe Arg Ser Tyr Val Phe Tyr Arg His Gln Gln Glu Gln Glu
            35   40   45

Ala Glu Gly Val Ala Ala Pro Ala Asp Pro Glu Met Val Thr Leu Pro
    50   55   60

Leu Gln Pro Ser Ser Thr Met Gly Gln Val Gly Arg Gln Leu Ala Ile
65   70   75   80

Ile Gly Asp Asp Ile Asn Arg Arg Tyr Asp Ser Glu Phe Gln Thr Met
                85   90   95

Leu Gln His Leu Gln Pro Thr Ala Glu Asn Ala Tyr Glu Tyr Phe Thr
            100   105   110

Lys Ile Ala Thr Ser Leu Phe Glu Ser Gly Asn Trp Gly Arg Val Val
            115   120   125

Ala Leu Leu Gly Phe Gly Tyr Arg Leu Ala Leu His Val Tyr Gln His
    130   135   140

Gly Leu Thr Gly Phe Leu Gly Gln Val Thr Arg Phe Val Val Asp Phe
145   150   155   160

Met Leu His His Cys Ile Ala Arg Trp Ile Ala Gln Arg Gly Gly Trp
                165   170   175

Val Ala Ala Leu Asn Leu Gly Asn Gly Pro Ile Leu Asn Val Leu Val
            180   185   190

Val Leu Gly Val Val Leu Leu Gly Gln Phe Val Val Arg Arg Phe Phe
            195   200   205

Lys Ser
210